Adaptive Subtype and Stage Inference for Alzheimer's Disease

Xinkai Wang[1,2] and Yonggang Shi[1,2]

[1] Stevens Neuroimaging and Informatics Institute, Keck School of Medicine, University of Southern California (USC), Los Angeles, CA 90033, USA
[2] Ming Hsieh Department of Electrical and Computer Engineering, Viterbi School of Engineering, University of Southern California (USC), Los Angeles, CA 90089, USA

Abstract. Subtype and Stage Inference (SuStaIn) is a useful Event-based Model for capturing both the temporal and the phenotypical patterns for any progressive disorders, which is essential for understanding the heterogeneous nature of such diseases. However, this model cannot capture subtypes with different progression rates with respect to predefined biomarkers with fixed events prior to inference. Therefore, we propose an adaptive algorithm for learning subtype-specific events while making subtype and stage inference. We use simulation to demonstrate the improvement with respect to various performance metrics. Finally, we provide snapshots of different levels of biomarker abnormality within different subtypes on Alzheimer's Disease (AD) data to demonstrate the effectiveness of our algorithm.

Keywords: Disease Heterogeneity · Subtype-Specific Events Discovery · Alzheimer's Disease.

1 Introduction

1.1 Background

Neurodegenerative diseases, such as Alzheimer's Disease (AD), are highly heterogeneous [9, 10, 14, 17], which means that individuals affected by the same medical conditions can have high variability in clinical, genetic, and environmental measurements, as well as treatment responses and prognosis. Better understandings of disease heterogeneity can improve diagnosis, treatment strategies, and patient outcomes with the assistance and advancement of personalized medicines.

1.2 Previous Model

There are two essential aspects of such heterogeneity: phenotypic heterogeneity (disease subtypes) [2, 12, 13] and temporal heterogeneity [8, 15] (disease stages for distinct subtypes). Most previous works concentrate on only one of the aspects, but the interactions between the two aspects might be missed. However, the Subtype and Stage Inference model [20] can effectively capture such relations

by dividing the disease population into distinct subgroups with different disease progression trajectories. Furthermore, based on the learned disease trajectories, and given a new individual, the model is able to predict which subgroup the individual most likely belongs to, as well as its most likely disease stage within that subgroup. As a result, this model can provide a more refined and precise stratification of patients based on both phenotypic and temporal aspects of a certain progressive disease and facilitates the development of precision medicine in clinical trials and healthcare.

1.3 Previous Model Mathematical Details

The Subtype and Stage Inference (SuStaIn) model is an extension of the event-based disease progression model [7], where each biomarker has a series of events as z-scores that measure the degree of abnormality for a given biomarker from the patient population. The model defines the events as hyperparameters before fitting the model, and the ultimate goal is to learn a permutation or ordering of the events both within and across biomarkers. The SuStaIn model learns a mixture of the continuous versions of the event-based disease progression model and implements piece-wise linear interpolation between events to form disease trajectories for each biomarker, where each trajectory measures the biomarker's z-score as a function of time.

Formally, define a set of subjects $j = 1, \ldots, J$ and each subject has biomarker measurements $i = 1, \ldots, I$. Then, the entire data is defined as $X = \{x_{ij} : i = 1, \ldots, I, j = 1, \ldots, J\}$. Define the time of occurrence of the events $\{t_{z^i_{k_i}} : i = 1, \ldots, I, k_i = 1, \ldots, R_i\}$, where each biomarker i has a total of R_i z-score events, and the events take values $\{z^i_{k_i} : i = 1, \ldots, I, k_i = 1, \ldots, R_i\}$. Furthermore, define $\{t_{z^i_{\max}} : i = 1, \ldots, I\}$ as the maximum z-score thresholds (z-max) for each biomarker. Let $R = \sum_{i=1}^{I} R_i$ be the total number of events or number of stages, and the goal of the model is to learn a bijection $\pi : \{1, \ldots, R\} \to \{1, \ldots, R\}$ that defines an ordering for

$$S = \{t_{z^1_1}, \ldots, t_{z^1_{R_1}}, t_{z^2_1}, \ldots, t_{z^2_{R_2}}, \ldots, t_{z^I_1}, \ldots, t_{z^I_{R_I}}\}.$$

The SuStaIn model tries to find disease subtypes $c = 1, \ldots, C$, and each subtype has a prior $f_c = P(S_c)$ and its data likelihood $P(X|S_c)$. The model finds $f_1, \ldots, f_C, S_1, \ldots, S_C$ that maximizes the following data likelihood:

$$P(X|M) = \sum_{c=1}^{C} f_c P(X|S_c),$$

where

$$P(X|S_c) = \prod_{j=1}^{J} \left[\int_0^{N+1} P(t) \prod_{i=1}^{I} P(x_{ij}|t) dt \right], \tag{1}$$

$$x_{ij}|t \sim N(g_i(t), \sigma_i),$$

and

$$g_i(t) = \begin{cases} \frac{z_1}{t_{z_1}} t, 0 < t \leq t_{z_1} \\ z_1 + \frac{z_2 - z_1}{t_{z_2} - t_{z_1}} \left(t - t_{z_1}\right), t_{z_1} < t \leq t_{z_2} \\ \vdots \\ z_{R-1} + \frac{z_R - z_{R-1}}{t_{z_R} - t_{z_{R-1}}} \left(t - t_{z_{R-1}}\right), t_{z_{R-1}} < t \leq t_{z_R} \\ z_R + \frac{z_{\max} - z_R}{t_{z_{\max}} - t_{z_R}} \left(t - t_{z_R}\right), t_{z_R} < t \leq t_{z_{\max}} \end{cases}$$

1.4 Motivation of Current Model

In general, the Z-Score SuStaIn model (abbreviated as SuStaIn) [1, 20] is able to capture both the phenotypical and temporal complexities by the following procedure:

1. Define the z-score events as hyperparameters of the model. First, define $z^i_{\max}$ for each biomarker as the maximum z-scored biomarker values from the entire data. Next, specify the number of z-score events for each biomarker as R_1. Then, the z-score events can be set as $\{z^i_r = \frac{z^i_{\max}}{R_1+1} r, i = 1, \ldots, I, r = 1, \ldots, R_1\}$, which splits $[0, z^i_{\max}]$ into equal intervals for interpolation and gives uniformly sampled information on the trajectories. This is the default option unless specified with some domain expertise.

2. Run the data X on SuStaIn model. The model will keep splitting all the data into distinct subgroups until further splitting yields a lower overall likelihood. Eventually, the model will output C permutation sequences, one sequence for each subtype. Then, combining the sequences and z-score events defined in step 1, one continuous trajectory for each biomarker in each subtype can be obtained by linear interpolation on the points $(t_{z^i_{k_i}}, z^i_{k_i})$ for fixed i.

3. Given the trajectories obtained from step 2, the model calculates the probability of each individual belonging to each subtype and each stage. Then the model will output the maximum likelihood subtype and stage for each individual.

However, the model assumes and defines a common set of z-score events for all subtypes beforehand. In particular, the maximum z-scores $(z^1_{\max}, z^2_{\max}, \ldots, z^I_{\max})$ is the same for all subtypes. Nevertheless, different subtypes may have different sets of maximum z-scores, and they may require different sets of z-score events defined accordingly from step 1. Alternatively speaking, the degree of abnormality of a single biomarker may be considered different for different disease subtypes. Previous works have demonstrated that AD patients have various disease progression rates [3–5, 11], which can be validated by the genetic burden measured by the cytosine-adenine-guanine (CAG) age product (CAP) score: fast progressing subjects tend to have higher genetic burden [19]. Another previous study has demonstrated that there are more than one possible distributions of

the rate of AD progression for any AD patient [18]. These imply that the same biomarker in different subtypes should have different rates of progression and may not reach the same abnormality level at late stages of the disease, and the model should allow the flexibility of reflecting such different rates by defining subtype-specific z-score events that are adaptive to each updated set of subtyped data during the model training process.

Hence, we propose the Subtype-Specific Events Discovery (SSED) Algorithm, which incorporates subtype-specific z-score event definitions. This approach allows for the characterization of possible disease stages tailored to each subpopulation from distinct subtypes, accommodating potentially inconsistent progression rates. SSED is particularly effective at identifying atypical subtype trajectories that exhibit significant differences in progression rates across ROIs. We relax the constraints for fixed z-score events definitions, enlarge the previous model space, which is the set of all possible outputs of the model, while maintaining the model time complexity.

2 Methods

2.1 Subtype-Specific Events Discovery (SSED) Algorithm Details

First, we run the SuStaIn model and obtain the maximum likelihood subtype for each subject $subInd$, and we initialize SSED with this subtype assignment. Then, the iterative procedure follows:

Fitting Step: we run SuStaIn separately ($RunSep$) on each subtyped datasets to obtain C event sequences. Combined with the z-score values of the events, we obtain C I-dimensional trajectories for each of the C subtypes.

Subtyping Step: given each of the subtype trajectories, we can calculate each subject's probability of belonging to each stage of each subtype ($CalProb$), which corresponds to a $J \times R \times C$ tensor P_s. Then, we add up all probabilities of stages for each subject and each subtype to form a $J \times C$ matrix P. Then, we add up all probabilities of subjects for each subtype to and normalize to form a C-dimensional vector f, which corresponds to the probability of each subtype $f_1, \ldots, f_C$. Denote the random variable St as the index of the subtype, and the random variable Sj as the index of the subject. According to the Bayes rule ($CalBayes$),

$$Q_{jc} := P(St = c|Sj = j) = \frac{P(Sj = j|St = c)P(St = c)}{\sum_{l=1}^{C} P(Sj = j|St = l)P(St = l)} = \frac{P_{jc}f_c}{\sum_{l=1}^{C} P_{jl}f_l}$$

Finally, we can find the updated subtype assignment indices $subIndNew$ by

$$c_j^* = \arg\max_c Q_{jc}.$$

As for finding new subtypes, SuStaIn uses divisive subtyping, which could potentially lead to problems because the new subtypes depend on previous divisions.

However, SSED will re-evaluate subtypes at different iterations to avoid the problem.

Next, we define several metrics for model evaluation. The first metric is data log likelihood, which is the same metric that SuStaIn uses [20]. However, we would like to define another performance metric that does not depend on the model . We leverage the silhouette score that measures clustering performance: the larger inter-cluster distance and smaller intra-cluster distance, the higher the score. However, optimal subtype assignments still have large distance between data from different stages, but data from different stages across subtypes follow the clustering structure. Since we do not have a more reasonable performance metric for the model at this time, we define the performance metric to be **the average stage-wise silhouette score** (ASSS), which calculates the average silhouette score of subsets of data that belong to the same stage across subtypes ($CalASSS$).

We will repeat the iterative procedure above until the ASSS has a small change between iterations, and the algorithm will output the events sequences seq and their corresponding z-score event values $zvals$ (points for interpolation of the subtype trajectories) for each subtype, maximum likelihood subtype $subNewInd$, and maximum likelihood stage $stageInd$ for each subject. In summary, the pseudocode for the algorithm is presented below.

Algorithm 1 Subtype-Specific Events Discovery (SSED)

procedure SSED($subInd, data, iter, score$)
  for c in $Unique(subInd)$ **do** ▷ Expectation Step
    $seq[c], zvals[c] \leftarrow RunSep(data[where(subInd == c)])$
  end for
  $Ps \leftarrow CalProb(seq)$ ▷ Maximization Step
  $P \leftarrow Sum(Ps, 1)$ ▷ Sum over all stages
  $f \leftarrow Sum(P, 0)/Sum(Sum(P, 0))$ ▷ Sum over all subjects
  $Q \leftarrow CalBayes(P, f)$
  $subIndNew \leftarrow Argmax(Q, 1)$ ▷ Maximize over all subtypes
  $stageInd, scoreNew \leftarrow CalASSS(subIndNew, P_s, data)$
  if $|scoreNew - score| \leq threshold$ **then** ▷ Terminate algorithm
    return $seq, zvals, subIndNew, stageInd$
  end if
  return SSED($subIndNew, data, iter + 1, scoreNew$) ▷ Next iteration
end procedure

Finally, since we do not know appropriate number of subtypes C, we will start from running 1 subtype model and keep increasing the number of subtypes and refit to SSED to see if the model likelihood keeps increasing. Whenever we observe a decrease in likelihood between fitting with C subtypes and $C + 1$ subtypes, we stop with the output with C subtypes.

2.2 Simulation

In this section, we will demonstrate that SSED is capable of capturing existing patterns. First, we set three ground truth piece-wise linear trajectories, and different colors represent different ground truth subtypes. Then, we extract the interpolation points $(t_{z^i_{k_i}}, z^i_{k_i})$ and add Gaussian noise to each of the points to form our simulation data. The lightness of the colors represent the ground truth stages of different subtypes: the darker of each point, the later stage it belongs to throughout the duration of disease progression. The fourth plot is the visualization of the SuStaIn output, including the learned trajectories, subtypes indicated by the color, and stages indicated by the lightness. We can see that SuStaIn can't distinguish two of the ground truth subtypes. However, we can continue to train the SSED using the 2 subtype assignments obtained from running SuStaIn. Afterwards, we try to fit 3 subtypes with SSED to see if the likelihood is higher than fitting 2 subtypes with SSED. As for the initialization of the third subtype, we randomly change 10% of the data to be the third subtype. At iteration 0, we can see that the subtypes are mixed together; however, at iteration 1, we can see that both the learned trajectories and subtypes shift towards the ground truth, and the algorithm eventually stops when the output almost exactly matches with the ground truth trajectories, subtypes, and stages (Fig. 1).

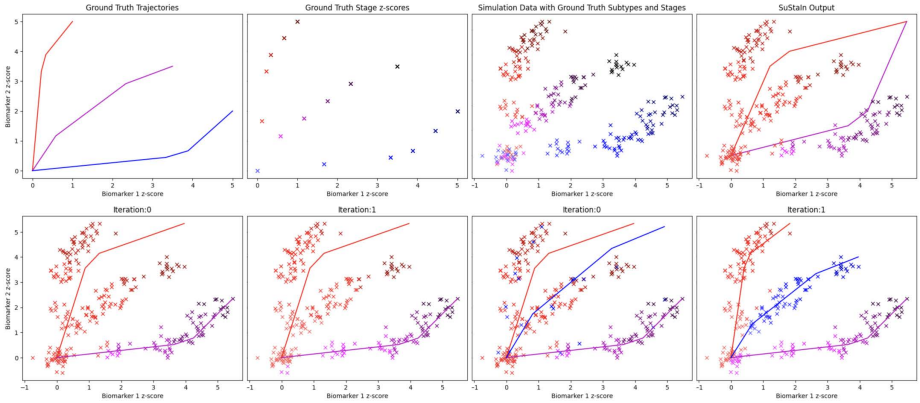

Fig. 1. Top Row: Ground truth trajectories and its simulation data; SuStaIn output. Bottom Row: SSED output for each iteration for 2 subtypes; SSED output for each iteration for 3 subtypes.

We quantitatively compare the performance in five metrics: total data likelihood according to equation (1); sum of Euclidean distances between ground truth interpolation points and output interpolation points at each stage of each subtype (Sum of Distance); subtype Adjusted Random Index (Subtype ARI), which is a similarity measure of two cluster assignments; stage Adjusted Ran-

dom Index (Stage ARI); and ASSS. From these metrics (Table 1), we see that running SSED with 3 subtypes has a higher performance than SuStaIn in all five metrics, showing that SSED is an improvement on the SuStaIn model.

Table 1. Performance Comparison between SuStaIn and SSED outputs.

	Log Likelihood	Sum of Distance	Subtype ARI	Stage ARI	ASSS
SuStaIn 2 Subtypes	-1044	30.26	.341	.424	.643
SSED 2 Subtypes	-880	8.933	.386	.356	.66
SSED 3 Subtypes	**-776**	**6.322**	**.593**	**.569**	**.701**

3 Results

3.1 Datasets and Data Preprocessing

We use 159 tau PET images from the Alzheimer's Disease Neuroimaging Initiative (ADNI) Study and 268 tau PET images from the Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease (A4) Study [16]. In total, we have 427 amyloid-positive-tau-positive (A+T+) cross-sectional PET images for training both SuStaIn and SSED models, where A+/A- labels are provided by ADNI and A4, and T+/T- labels are based on whether the peak SUVR value of the image exceeds 1.5. All images are first preprocessed by Freesurfer [6] for volume parcellation using the Desikan-Killiany atlas, averaged across frames and registered to T1 space. Then, the images are corrected for partial volume effect and intensity-normalized with the inferior cerebellar gray matter reference region to get SUVR images. Then, the SUVR images are projected onto individual cortical surfaces and the *fsavergae_rh* cortical template [21]. Then we extract the average SUVR values across each of the five ROIs as biomarkers for each subject: Temporal, Frontal, Parietal, Occipital, and Medial Temporal Lobe. Finally, we z-score all the data with respect to the cognitively normal (CN A-T-) population by its mean and standard deviation.

3.2 Results on ADNI and A4 Datasets

Table 2. Performance Comparison between SuStaIn and SSED outputs.

	ADNI Likelihood	ADNI ASSS	A4 Likelihood	A4 ASSS	ADNI&A4 ARI
SuStaIn	-1646	.276	-2159	.209	.155
SSED	**-1401**	**.343**	**-1931**	**.294**	**.561**

We run SSED on ADNI and A4 data separately. We choose to fit 3 subtypes, which agrees with the number of subtypes detected in previous work on SuStaIn

model [20]. Since we don't have ground truth subtypes and stages, we will stay with the remaining measures: model log likelihood and ASSS. In order to further validate the improvement on SSED, since ADNI and A4 data is similarly distributed, we use ARI to measure the consistency of output subtypes between ADNI and A4. Specifically, we use the trained SSED model on ADNI to predict subtype assignments of all the A4 data and compare their similarity with subtype assignments of all the A4 data directly from the trained SSED model on A4. We set the number of events defined for each biomarker to be 2 for efficiency, and we have 5 biomarkers, which gives 10 stages in total for each subtype. We can see from the table (Table 2) that SSED has a higher performance over SuStaIn in all the metrics defined.

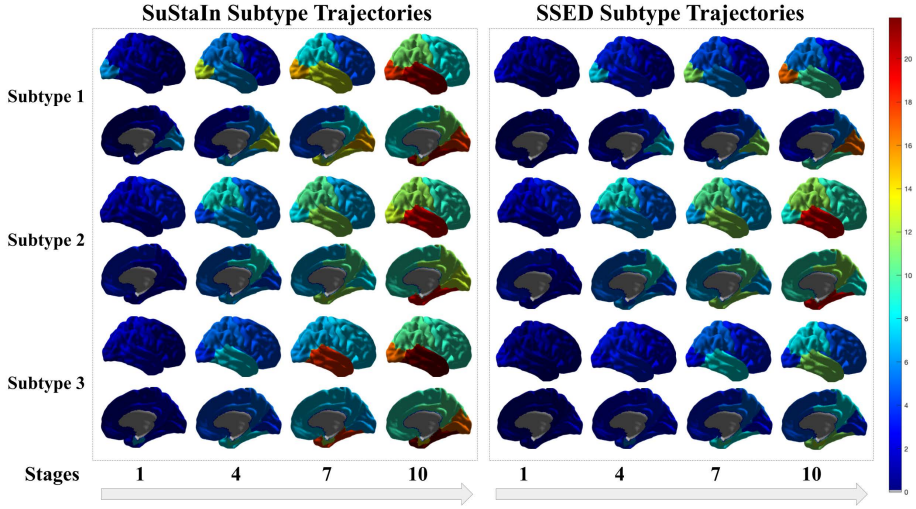

Fig. 2. SuStaIn and SSED output subtype sequences for stages 1, 4, 7, and 10.

Finally, we can visualize the subtype patterns of ADNI data throughout stages (Fig. 2). The color scale represents the z-scored SUVR values. We compare the SuStaIn and SSED outputs of the snapshots of subtypes trajectories at stages 1, 4, 7, and 10. The subtype patterns are similar for both outputs. However, SuStaIn subtypes are forced to converge to similar patterns due to the constraint of a single set of definition of events for distinct subtypes, but SSED subtypes can decorrelate biomarkers and are flexible to have biomarker values with more significant difference between not only ROIs but also subtypes at late stages.

4 Conclusions

In summary, SuStaIn is able to capture temporal and phenotypical patterns for progressive disease, and we propose an algorithm for improving its performance

with adaptive subtype-specific events, which is a more appropriate and reasonable assumption of the model. We demonstrate the effectiveness of our algorithm with simulation data, and we can visualize how our algorithm can generate more distinct subtype patterns with ADNI data and capture more heterogeneous information from the same data.

Acknowledgments. This work is supported by the National Institute of Health (NIH) under grants R01EB022744, RF1AG077578, RF1AG064584, R21AG064776, R01AG062007, U19AG078109, and P41EB015922.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Aksman, L.M., Wijeratne, P.A., Oxtoby, N.P., Eshaghi, A., Shand, C., Altmann, A., Alexander, D.C., Young, A.L.: pysustain: a python implementation of the subtype and stage inference algorithm. SoftwareX **16**, 100811 (2021)
2. Bird, T., Sumi, S., Nemens, E., Nochlin, D., Schellenberg, G., Lampe, T., Sadovnick, A., Chui, H., Miner, G., Tinklenberg, J.: Phenotypic heterogeneity in familial alzheimer's disease: a study of 24 kindreds. Annals of neurology **25**(1), 12–25 (1989)
3. Byun, M.S., Kim, S.E., Park, J., Yi, D., Choe, Y.M., Sohn, B.K., Choi, H.J., Baek, H., Han, J.Y., Woo, J.I., et al.: Heterogeneity of regional brain atrophy patterns associated with distinct progression rates in alzheimer's disease. PLoS One **10**(11), e0142756 (2015)
4. Doody, R.S., Pavlik, V., Massman, P., Rountree, S., Darby, E., Chan, W.: Predicting progression of alzheimer's disease. Alzheimer's research & therapy **2**, 1–9 (2010)
5. Duara, R., Barker, W.: Heterogeneity in alzheimer's disease diagnosis and progression rates: implications for therapeutic trials. Neurotherapeutics **19**(1), 8–25 (2023)
6. Fischl, B., Van Der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., et al.: Automatically parcellating the human cerebral cortex. Cerebral cortex **14**(1), 11–22 (2004)
7. Fonteijn, H.M., Modat, M., Clarkson, M.J., Barnes, J., Lehmann, M., Hobbs, N.Z., Scahill, R.I., Tabrizi, S.J., Ourselin, S., Fox, N.C., et al.: An event-based model for disease progression and its application in familial alzheimer's disease and huntington's disease. NeuroImage **60**(3), 1880–1889 (2012)
8. Jack, C.R., Knopman, D.S., Jagust, W.J., Shaw, L.M., Aisen, P.S., Weiner, M.W., Petersen, R.C., Trojanowski, J.Q.: Hypothetical model of dynamic biomarkers of the alzheimer's pathological cascade. The Lancet Neurology **9**(1), 119–128 (2010)
9. Jellinger, K.A.: Recent update on the heterogeneity of the alzheimer's disease spectrum. Journal of Neural Transmission **129**(1), 1–24 (2022)
10. Lam, B., Masellis, M., Freedman, M., Stuss, D.T., Black, S.E.: Clinical, imaging, and pathological heterogeneity of the alzheimer's disease syndrome. Alzheimer's research & therapy **5**(1), 1–14 (2013)

11. Lucca, U., Comelli, M., Tettamanti, M., Tiraboschi, P., Spagnoli, A.: Rate of progression and prognostic factors in alzheimer's disease: a prospective study. Journal of the American Geriatrics Society **41**(1), 45–49 (1993)
12. Murray, M.E., Graff-Radford, N.R., Ross, O.A., Petersen, R.C., Duara, R., Dickson, D.W.: Neuropathologically defined subtypes of alzheimer's disease with distinct clinical characteristics: a retrospective study. The Lancet Neurology **10**(9), 785–796 (2011)
13. Ryan, J., Fransquet, P., Wrigglesworth, J., Lacaze, P.: Phenotypic heterogeneity in dementia: a challenge for epidemiology and biomarker studies. Frontiers in public health **6**, 181 (2018)
14. Schneider, J.A., Arvanitakis, Z., Bang, W., Bennett, D.A.: Mixed brain pathologies account for most dementia cases in community-dwelling older persons. Neurology **69**(24), 2197–2204 (2007)
15. Sperling, R.A., Aisen, P.S., Beckett, L.A., Bennett, D.A., Craft, S., Fagan, A.M., Iwatsubo, T., Jack Jr, C.R., Kaye, J., Montine, T.J., et al.: Toward defining the preclinical stages of alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. Alzheimer's & dementia **7**(3), 280–292 (2011)
16. Sperling, R.A., Rentz, D.M., Johnson, K.A., Karlawish, J., Donohue, M., Salmon, D.P., Aisen, P.: The a4 study: stopping ad before symptoms begin? Science translational medicine **6**(228), 228fs13–228fs13 (2014)
17. Sun, N., Mormino, E.C., Chen, J., Sabuncu, M.R., Yeo, B.T., Initiative, A.D.N., et al.: Multi-modal latent factor exploration of atrophy, cognitive and tau heterogeneity in alzheimer's disease. Neuroimage **201**, 116043 (2019)
18. Thalhauser, C.J., Komarova, N.L.: Alzheimer's disease: rapid and slow progression. Journal of the Royal Society Interface **9**(66), 119–126 (2012)
19. Wijeratne, P.A., Eshaghi, A., Scotton, W.J., Kohli, M., Aksman, L., Oxtoby, N.P., Pustina, D., Warner, J.H., Paulsen, J.S., Scahill, R.I., et al.: The temporal event-based model: Learning event timelines in progressive diseases. Imaging Neuroscience **1**, 1–19 (2023)
20. Young, A.L., Marinescu, R.V., Oxtoby, N.P., Bocchetta, M., Yong, K., Firth, N.C., Cash, D.M., Thomas, D.L., Dick, K.M., Cardoso, J., et al.: Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with subtype and stage inference. Nature communications **9**(1), 4273 (2018)
21. Yue, J., Shi, Y.: Uncovering heterogeneity in alzheimer's disease from graphical modeling of the tau spatiotemporal topography. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 262–271. Springer (2023)